ABSTRACT

A distributed automatic speech recognition and text to speech system has components separated by a network. The components may be a voice interface device with jitter buffer, a voice browser and a text to speech engine. A barge-in detection feature may reside in any one of these components, to implement kill on barge in. The system, softwares and methods of the invention operate to flush the jitter buffer of the voice interface device when a barge-in is detected. Any packets that had been received are therefore not played out.